

Analysis of a Coded, M-ary Orthogonal Input Optical Channel With Random-Gain Photomultiplier Detection

P. J. Lee

Communications Systems Research Section

Performance of two coding systems is analyzed in this study for a noisy optical channel with $M(= 2^L)$ -ary orthogonal signaling and random-gain photomultiplier detection. The considered coding systems are the Reed-Solomon (RS) coding with error-only correction decoding and the interleaved binary convolutional system with soft decision Viterbi decoding. The required average number of received signal photons per information bit, \bar{N}_b , for a desired bit error rate of 10^{-6} is found for a set of commonly used parameters and with a high background noise level. We find that the interleaved binary convolutional coding system is preferable to the RS coding system in performance-complexity tradeoffs.

I. Introduction

The utilization of optical frequencies for deep-space communications has been considered as an attractive alternative to microwave frequencies, because of the advantages of smaller apparatus sizes and wider bandwidths. $M(= 2^L)$ -ary orthogonal signaling is the most popular modulation scheme for energy-efficient optical communications when a large bandwidth-bit time product is available. Such a signaling is implemented by multiple-pulse-position modulation and/or multiple-color (frequency) modulation (Refs. 1 and 2). Additionally, direct detection employing a photomultiplier is often used for reception.

For further energy efficiency, coding over such a channel has been studied. When no background noise and ideal detection are assumed, the above channel can be modeled as an

M -ary input, $(M + 1)$ -ary output erasure-only channel (MEC). For this simplest channel, Reed-Solomon (RS) coding with erasure-only correction decoding was studied in Ref. 3. In Ref. 4, it was recognized that there is an equivalence between the MEC and the L -parallel correlated binary input, ternary output, erasure-only channels (BECs). Also in Ref. 4, the use of L -parallel, independent binary convolutional coding systems with Viterbi decoding were suggested; these systems exhibited performance favorable to those with similar complexity described in Ref. 3. For the channel with random-gain photomultiplier tube (PMT) reception, an RS coding system with combined error-and-erasure correction decoding was studied in Ref. 5, but with a negligibly low background noise level.

In this article, we consider the M -ary orthogonal input optical channel with random-gain PMT detection and with arbitrary

trary background noise levels, as described in the next section. The performance of two coding systems over this channel is then studied. In Section III, the performance analysis of an RS coding system with error-only correction decoding is given. Then by extending the channel decomposition idea of Ref. 4, we propose a method of using interleaved, short-constraint-length binary convolutional codes with soft decision Viterbi decoding for this noisy optical channel, and analyze its performance. A numerical example with a high background noise level is given. We conclude that, as in the case of the noiseless ideal optical channel, the interleaved binary convolutional coding system is preferable to the RS coding system for the noisy optical channel with nonideal photon counting, which results from realistic PMTs.

II. Noisy M-ary Orthogonal Input Optical Channel With a Random-Gain PMT Receiver

Let an optical transmitter emit an optical pulse in one pulse position among M_0 nonoverlapping pulse positions with one color among M_1 nonoverlapping colors at each symbol time T_s . Suppose the receiver can separate the pulse positions and the colors. Then we have an $M(=M_0 \times M_1)$ -ary input optical channel. Let X be the channel input symbol random variable, which takes its value from the M -ary channel input symbol alphabet $\mathcal{M} = \{0, \dots, m, \dots, M-1\}$. For each T_s , the receiver is to produce an M -dimensional observation vector $\mathbf{y} = (y_0, \dots, y_m, \dots, y_{M-1})$ where each y_m is the output at the detector consisting of a PMT and an integrate-and-dump filter for the assigned color-time slot. A model for the detector is shown in Fig. 1, where $T_c (= T_s/M_0)$ is the slot time and R is the PMT anode load resistance. Note that we need such a PMT detector for each M_1 color slot and the color splitter. We assume symmetries and independences for every time and color slot so that the resulting channel becomes an M -ary orthogonal input, symmetric output, memoryless channel. We restrict our attention to $M = 2^L$ cases with integer L .

A PMT is characterized by its average gain A and the number of dynodes ν . The variance of the PMT gain is $2AB$ where $2B = (A-1)/(A^{1/\nu}-1)$. In Ref. 6, a Markov diffusion approximation to the statistics of PMT gain was given. Including the effect of thermal noise, the statistics of the detector output voltage random variable Y , conditioned on the number of input photons a , is given by (Ref. 6)

$$\begin{aligned} H(y|a) &= \Pr \{Y \leq y|a\} \\ &= C \cdot Q(-y/s) + \int_0^\infty Q(-(y-cz)/s) \cdot h(z|a) \cdot dz \end{aligned}$$

where

$$\begin{aligned} h(z|a) &= e^{-a} \cdot \sum_{n=1}^{\infty} \frac{1}{n!} (a \cdot e^{-A/B})^n \\ &\quad \cdot \sum_{k=1}^{\infty} \frac{(nA/B^2)^k \cdot z^{k-1}}{k!(k-1)!} e^{-z/B} \\ Q(y) &\equiv \int_y^\infty \exp(-t^2/2) \cdot dt/\sqrt{2\pi}, \end{aligned}$$

$C = \exp[-a\{1 - \exp(-A/B)\}]$, s^2 is the variance of thermal noise ($= N_0/T_c$, N_0 is the one-sided thermal noise spectral density), and $c = eR/T_c$ (e is the electron charge).

Let \bar{N}_n be the average number of noise photons per slot and \bar{N}_s be the average number of received signal photons per M -ary symbol. Recall that there is only one active slot among M possible slots. Let $G(y)$ be the cumulative distribution function for the active slot and $F(y)$ be that for a nonactive slot. Then we have

$$G(y) = H(y|a = \bar{N}_s + \bar{N}_n) \text{ and } F(y) = H(y|a = \bar{N}_n)$$

and the channel transition probabilities for each $m \in \mathcal{M}$ can be represented as

$$\Pr \{Y \leq y | X = m\} = G(y_m) \cdot \prod_{m' \neq m} F(y_{m'})$$

III. Reed-Solomon Coding System With Error-Only Correction Decoding

For an M -ary orthogonal input, symmetric output channel, one of the best coding systems is the (N, K) RS coding system whose symbol alphabet size is M , where K and N are the numbers of information and channel symbols in an RS codeword (Ref. 7). The code rate r is K/N [information symbols/channel symbol], and N is usually chosen to be $M-1$ or M . When the decoder is capable of correcting errors only, a hard decision (i.e., M -ary symbol decision) must be performed in front of the decoder. By assuming equiprobable channel input symbols, it is easy to show that the maximum likelihood (ML) symbol decision rule is optimum due to the symmetries of the channel. Since $dG(y)/dF(y)$ is a monotonically increasing function of y for this optical channel, the decision rule simplifies to, "declare m as the transmitted symbol, if the m -th observation value y_m is the largest." The corresponding decision error probability, p , is given by

$$p = (M-1) \cdot \int_{-\infty}^{\infty} G(y) \cdot F^{M-2}(y) \cdot dF(y)$$

The resulting RS decoded bit-error rate (BER) is given by (Ref. 7)

$$BER = \frac{M}{2(M-1)} \cdot \sum_{i=t+1}^N \binom{N}{i} \cdot p^i \cdot (1-p)^{N-i} \cdot \frac{i}{N}$$

where t ($=$ integer part of $(N-K)/2$) is the number of correctable error symbols in an RS codeword.

IV. Binary Convolutional Coding System With Soft Decision Viterbi Decoding

For binary-input symmetric-output channels, a short-constraint-length binary convolutional coding system with soft decision Viterbi decoding is one of the most practical and powerful coding systems available. Let $U \in \{0, 1\}$ be the binary channel input symbol and $\mu(y, U)$ is the metric to be used by the decoder for the channel input symbol U with a given channel output observation vector y . Let D be the Chernhoff bound on the probability that the metric for a transmitted symbol is smaller than that for a nontransmitted symbol. Then we have (Ref. 8)

$$D = \min_{w \geq 0} D(w)$$

$$D(w) = E[\exp \{w \cdot [\mu(Y, U = 1) - \mu(Y, U = 0)]\} | U = 0]$$

where the operator E is the expectation over the observation random vector Y . Notice that D becomes the Bhattacharyya bound when the metric is an ML metric. For a given binary convolutional encoder having a transfer function $T(Z, I)$, the decoded BER at the output of the associated Viterbi decoder using the metric μ is well bounded by (Ref. 8)

$$BER \leq 1/2 \cdot \left. \frac{\partial}{\partial I} T(Z, I) \right|_{Z=D, I=1}$$

Recall that $M = 2^L$ and each one of the M -ary symbols is assumed to have a one-to-one correspondence with a set of L bits. An example of such a mapping is shown in Table 1 for the $M = 8$ ($L = 3$) case. Suppose we have an MEC with erasure probability of q . That is, the transmitted symbol is either correctly received with probability $1 - q$ or is erased with probability q . Hence, once the receiver has an unerased symbol, all the corresponding L bits are recovered correctly. Alternately, when the receiver has an erased symbol, all the L bits are also erased. Accordingly, the MEC is equivalent to L -parallel binary

erasure-only channels (BEC), each with the same erasure probability q . Such component channels are fully correlated, since the erasure events always occur simultaneously for all of L -component channels. Massey, in Ref. 4, noticed this equivalence and suggested use of the binary convolutional coding system with Viterbi decoding for each component binary input channel independently (i.e., by ignoring the correlation). Note that, instead of using L -parallel binary coding systems, one may use one binary coding system with interleaving (Ref. 4). Also, the binary convolutional coding system was shown to be preferable to RS coding system in performance-complexity tradeoffs, for the ideal optical channel with no background noise.

Now we shall extend the Massey's channel decomposition idea for the M -ary optical channel to the case of nonideal detection and with background noise as described in Section II. Let \mathcal{I}_ℓ (\mathcal{J}_ℓ) represent the set of all M -ary channel input symbols having a 0 (1) in the ℓ -th component of their binary representations (respectively). For the example shown in Table 1, $\mathcal{I}_2 = \{0, 1, 4, 5\}$ and $\mathcal{J}_3 = \{1, 3, 5, 7\}$. Let i_ℓ and j_ℓ be the largest value among the channel output observation values in the set \mathcal{I}_ℓ and \mathcal{J}_ℓ . For the example set of observation values in Table 1, $i_2 = \max(3.11, 2.72, 6.67, 4.06) = 6.67$ and $j_3 = 4.06$. We may use this i_ℓ and j_ℓ directly for the metric for the ℓ -th component channel input binary symbol U_ℓ , i.e., $\mu(y, U_\ell = 0) = i_\ell$ and $\mu(y, U_\ell = 1) = j_\ell$. This is not an ML metric for such a binary input coding channel, but is a good choice based on the complexity consideration. It is easy to show that the corresponding Chernhoff bounds are given, for each $\ell = 1, 2, \dots, L$, by

$$D_\ell = D = \min_{w \geq 0} D(w)$$

$$D(w) = \left[(M/2) \cdot \int_{-\infty}^{\infty} \exp(wy) \cdot F^{M/2-1}(y) \cdot dF(y) \right] \cdot \left[\int_{-\infty}^{\infty} \exp(-wy) \left\{ \left(\frac{M}{2} - 1 \right) \cdot G(y) \cdot F^{M/2-2}(y) \cdot dF(y) + F^{M/2-1}(y) \cdot dG(y) \right\} \right]$$

where we have used the following:

$$\begin{aligned} \Pr \{I_\ell \leq y | U_\ell = 0\} &= \Pr \{J_\ell \leq y | U_\ell = 1\} \\ &= G(y) \cdot F^{M/2-1}(y) \end{aligned}$$

$$\Pr \{I_\ell \leq y | U_\ell = 1\} = \Pr \{J_\ell \leq y | U_\ell = 0\} = F^{M/2}(y)$$

where I_q and J_q are the random variables corresponding to i_q and j_q , respectively.

V. Numerical Example, Discussion, and Conclusions

The coding systems described and analyzed in the previous sections are compared for the specific example of an optical channel with random-gain PMT reception and a very high background noise level of 10^8 noise photons per second. Such a background level is expected when a spacecraft is in front of a bright planet. All the other parameters are chosen to be the same as those in Refs. 5 and 6: the slot time $T_c = 10^{-7}$ s, average PMT gain $A = 10^6$, the number of PMT dynodes $\nu = 11$, the PMT anode load resistance $R = 50 \Omega$, and the one-sided thermal noise power spectral density $N_0 = 1.156 \times 10^{-17}$ V/Hz.

With these parameters and the equations in the previous sections, we find the required average number of received signal photons per symbol, \bar{N}_s , for a desired BER of 10^{-6} and for some values of channel symbol alphabet size M and code rate r . For the RS coding system, $5 \leq L \leq 8$ and all possible code rates $r = K/N$ were considered with $N = M$. For the interleaved binary coding system, we consider $4 \leq L \leq 10$ and $r = 1/3, 1/2, 2/3, 3/4$, and $4/5$. For the latter system, there is another choice of code parameter, which is the number of states, S , in the Viterbi decoding. For a regular rate $1/n$ convolutional code with constraint length k , S is given by 2^{k-1} . For this numerical example, we restricted S to be 64 (or equivalently, $k = 7$), and used the (7, 1/2) and (7, 1/3) codes (found in Ref. 9) and the $r = 2/3, 3/4$, and $4/5$ punctured codes derived from the (7, 1/2) codes (found in Ref. 10). For such punctured codes, a decoder for the (7, 1/2) code can be used directly with minor modifications on metric forming (Ref. 10).

The quantity of interest for energy-efficient optical communication is the average number of received photons per information bit, \bar{N}_b , rather than \bar{N}_s . Note that $\bar{N}_b = 1/\rho$ where ρ is the commonly used parameter (Ref. 5) for photon efficiency in [bits/photon]. \bar{N}_b is related to \bar{N}_s by $\bar{N}_b = \bar{N}_s/(rL)$. The curves of required values of \bar{N}_b for $\text{BER} \leq 10^{-6}$ versus code rate r are plotted in Fig. 2. The solid lines correspond

to the RS coding system performances while the broken lines refer to the interleaved binary convolutional coding system performances.

From Fig. 2 we see that RS coding systems perform better for larger values of L (>6), while the interleaved binary 64-state convolutional coding systems perform better for smaller values of L (<6). However, notice that the RS coding system complexity grows exponentially as L increases, while the complexity of the binary coding system remains almost the same for any value of L . Hence, for a fair comparison between the two coding systems, one must consider the system complexity as well as the performance. For a given channel input alphabet size $M = 64$ ($L = 6$), the two coding systems show almost identical performance, but the interleaved 64-state binary Viterbi decoding system is considered to be less complex than the 6-bit RS decoding system.

For better performance with the interleaved binary coding system, one may increase the system complexity by employing longer constraint-length codes. The interleaved binary 256-state ($k = 9$) Viterbi decoded convolutional coding system performs almost the same as the 7-bit RS coding system for the noisy 128-ary optical channel and requires less complexity. From these performance-complexity considerations, we conclude that the interleaved binary convolutional coding system is preferable to the RS coding system for the noisy nonideal channel, similar to the results for the noiseless ideal optical channel considered in Ref. 4.

One may consider some variations of the two coding systems. For the RS coding system, instead of error-only correction decoding, one can consider a combined error-and-erasure correction decoding. This modification can improve performance, but the amount of improvement is usually very small except for the cases of very low background noise level, even with properly designed $(M + 1)$ -ary decision rules (see Ref. 7). The metric given in Section IV is the simplest unquantized metric for the component channel. Hence, one may consider better metrics that more closely approximate ML decoding metric. Furthermore, one can consider concatenated coding systems employing the interleaved convolutional coding as an inner coding. Hence, there is more flexibility and more room for performance improvement with the interleaved convolutional coding system.

Acknowledgment

The author would like to thank Prof. J. K. Omura of U.C.L.A. and Dr. J. R. Lesh of JPL for their encouragement and helpful discussions.

References

1. Pierce, J. R., "Optical Channels: Practical Limits With Photon Counting," *IEEE Trans. Comm.*, Vol. COM-16, pp. 1819-1821, Dec. 1978.
2. Davidson, F., "Direct-Detection Optical Communication with Color Coded Pulse Position Modulation Signaling," submitted to *IEEE Trans. Commun.*
3. McEliece, R. J., "Practical Codes for Photon Communication," *IEEE Trans. Infor. Theory*, Vol. IT-27, pp. 393-408, July 1981.
4. Massey, J. L., "Capacity, Cutoff Rate, and Coding for a Direct-Detection Optical Channel," *IEEE Trans. Commun.*, Vol. COM-29, pp. 1615-1621, Nov. 1981.
5. Lesh, J. R., J. Katz, H. H. Tan, and D. Zwillinger, "2.5 Bits/Detected Photon Demonstration Program: Analysis and Phase I Results," *TDA Progress Report 42-66*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 115-132, Dec. 1981.
6. Tan, H. H., "A Statistical Model of the Photomultiplier Gain Process With Applications to Optical Pulse Detection," *TDA Progress Report 42-68*, Jet Propulsion Laboratory, Pasadena, Calif., pp. 55-67, April 1982.
7. Berlekamp, E. R., "The Technology of Error-Correcting Codes," *Proceedings of IEEE*, Vol. 68, pp. 564-593, May 1980.
8. Omura, J. K., and M. K. Simon, *Modulation/Demodulation Techniques for Satellite Communications, Part IV. Appendices*, JPL Publication 81-73, Jet Propulsion Laboratory, Pasadena, Calif., Nov. 1981.
9. Odenwalder, J. P., *Optimal Decoding of Convolutional Codes*, Ph.D. Thesis, School of Engineering and Applied Sciences, U.C.L.A., Los Angeles, Calif., 1970.
10. Yasuda, Y., K. Kashiki, and Y. Hirata, "High-Rate Punctured Convolutional Codes for Soft Decision Viterbi Decoding," *IEEE Trans. Commun.*, Vol. COM-32, pp. 315-319, March 1984.

Table 1. An example of M-ary symbol to L binary symbols mapping and an example set of observation values

m	$u_1 u_2 u_3$	y_m
0	0 0 0	3.11
1	0 0 1	2.72
2	0 1 0	1.53
3	0 1 1	1.18
4	1 0 0	6.67
5	1 0 1	4.06
6	1 1 0	1.74
7	1 1 1	0.98

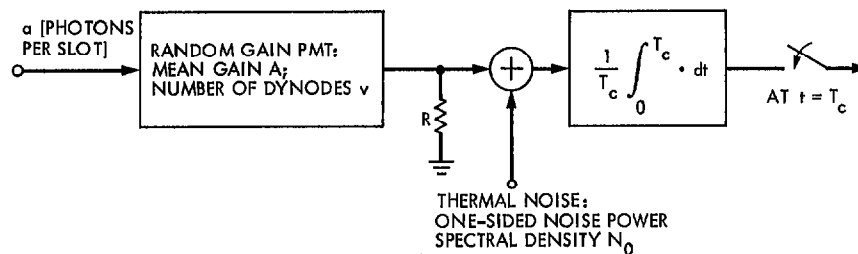


Fig. 1. A model for an optical receiver employing a random-gain photomultiplier tube

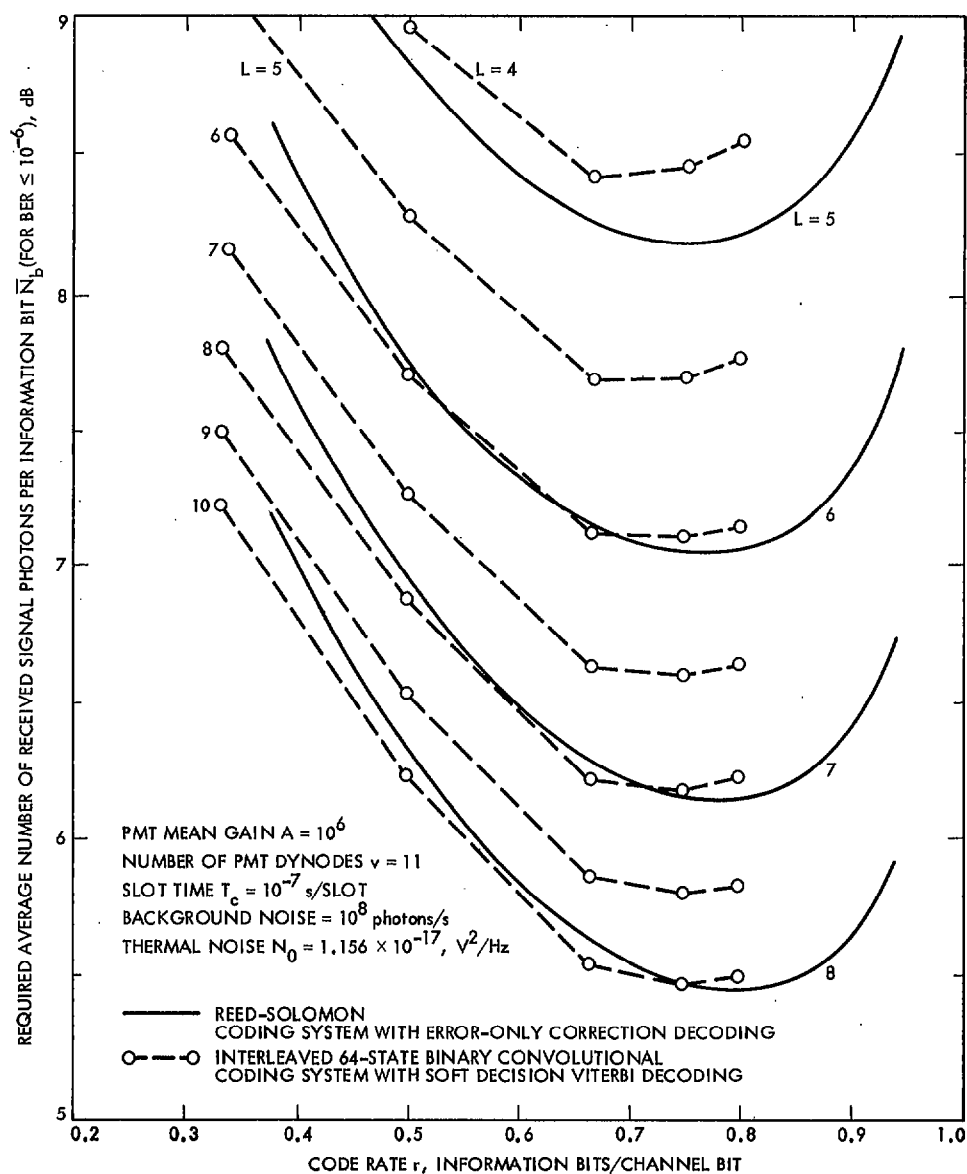


Fig. 2. Performances of coding systems for $M(= 2^L)$ -ary orthogonal input noisy optical channel